DOCUMENT RESUME

ED 262 076 TM 850 553

AUTHOR Phillips, Gary W.; Grodsky, Milton
TITLE Testing Piaget's Theory of Probability Concept Development: A Bayesian Approach Using the Theory of Signal Detection.
PUB DATE Apr 85
NOTE 44p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Age Differences; Bayesian Statistics; Cognitive Development; *Cognitive Measurement; Concept Formation; *Decision Making; Elementary Education; Females; Higher Education; Mathematical Models; *Nonverbal Tests; Preschool Education; *Probability
IDENTIFIERS Peabody Picture Vocabulary Test; Piagetian Theory; *Signal Detection Theory

ABSTRACT

The decision making processes of children in a probabilistic environment were studied within the context of the theory of signal detection (TSD). The relationship between the age of the child and his ability to revise decision criteria was examined, as well as the appropriateness of TSD measures and methodology with children. It was hypothesized that the decision criteria of older children would be more optimal than the younger children. Twenty-four females with normal I.Q.'s at three age levels (kindergarten, elementary school, and college) were administered a visual signal detection task. The prior probability of the signal was manipulated for half the subjects, and the value of the signal was manipulated for the other half. The results did not support the experimental hypothesis, since there was no age difference in decision criteria change. In addition, the results indicated that children as young as 4 to 5 years of age did operate according to TSD given the conditions of the model. Therefore, the TSD paradigm is appropriate for studying probabilistic thinking in children. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262076

TESTING PIAGETS THEORY OF PROBABILITY CONCEPT DEVELOPMENT: A BAYESIAN APPROACH USING THE THEORY OF SIGNAL DETECTION

Gary W. Phillips
Maryland State Department of Education
and
Milton Grodsky
University of Maryland
1985

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Phillips, Gary W.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 850 553

Paper presented at the Annual Convention of the American Education Research Association, Chicago, April 1985.

TESTING PIAGETS THEORY OF PROBABILITY CONCEPT DEVELOPMENT: A BAYESIAN APPROACH USING THE THEORY OF SIGNAL DETECTION

ABSTRACT

The study was concerned with the decision making processes of children in a probabilistic environment within the context of the theory of signal detection. More specifically, the research dealt with: (1) the relationship between the age of the child and his ability to revise decision criteria; and (2) the appropriateness of TSD measures and methodology for childhood populations. The experimental hypothesis was that the decision criteria of older children would be more optimal than the younger children.

24 female subjects with normal I.Q.'s at three age levels (day care, elementary school, and college) were utilized in a visual detection task. The signal detection task consisted of two overlapping triangular distributions of bar heights with equal variance. At each age level the prior probability of the signal was manipulated for 4 subjects, and the value of the signal was manipulated for the other 4 subjects.

The result did not support the experimental hypothesis in that there was no age difference in decision criteria change. This result was discussed in terms of the literature on probabilistic thinking in children.

In addition, the results indicated that children as young as 4 to 5 years of age do operate according to TSD given the conditions of the model. Therefore, the TSD paradigm provides appropriate measures and methodology for studying probabilistic thinking in children.

TESTING PIAGETS THEORY OF PROBABILITY CONCEPT DEVELOPMENT: A BAYESIAN APPROACH USING THE THEORY OF SIGNAL DETECTION

INTRODUCTION

The purpose of this research is to study the decision making processes of children in a probabilistic environment within the context of the Theory of Signal Detectability (TSD). At the present time there is no literature of decision making in children using a TSD paradigm. Therefore, this research is seen not only as an initial attempt at studying children's decisions within a TSD framework, but equally important, it is an attempt to determine the appropriateness of TSD measures and methodology for childhood populations.

The first part of this introduction will be concerned with a review of the literature on decision making in children in probabilistic environments. In addition to providing background information, this review will be used as a guide in the selection of the age ranges in the current study. Following the literature review, a brief description of TSD will be presented.

Much of the early work on decision making in probabilistic environments (with children) involved two-choice probability learning experiments. The main impetus to this research was provided by Jean Piaget (1960). Piaget argued that the concept of probability developed in two stages. During the Preoperational stage (3-6 years of age) the child does not distinguish possibility from necessity, does not respond to unequal proportions, and therefore has no concept of probability (Piaget, 1960). In the Concrete-Operational stage (7-10 years of age) the child first develops the rudiments of the probability concept, which are finally mastered in the stage of Formal-Operations (Piaget and Inhelder, 1951). During this latter stage

the child develops abstract concepts of combinations, permutations, and proportions.

In one study, for example, Piaget (1950) used 14 boys and girls between the ages of 5-12 to perform a probability task. The task involved an assortment of tokens with various colors placed in an opaque bag. An identical assortment of colored tokens were placed in a row in front of the children. Each child individually was asked which color would be most likely get if he reached into the bag without looking. Once withdrawn from the bag the tokens remained on the table in front of the child. Using his clinical interview assessment technique, Piaget concluded: (1) the younger children (between 5-7) did not make decisions based on the proportion of colors. Rather, they used irrelevant criteria such as the order of the tokens on the table, or personal preference for certain colors; (2) children between the ages of 8-10 based their decisions on the relative proportions of colors, but did not take into account sampling without replacement; (3) the older children recognized the changes in proportions, and therefore probabilities, due to lack of replacement.

The general trend in the current American research supports Piaget's notion that the concept of probability occurs in developmental stages. However, the onset of the concept appears to be much earlier than Piaget's theories and research indicate. For example, Siegel and Andrews (1962) studies the behavior of 24 four-year-old boys in a probability task, using amount of reinforcement as an independent variable. Two identical containers were placed before each subject, and the child was required to guess in which container was an object. In container A the object was present in 75% of the trials, while container B housed the object 25% of the trials (in a randomized series). In the low reinforcement condition a valueless button

was in the container. In the high reinforcement condition a valued prize (e.g., candy, toys, crayons, etc.) was in the container. Under the low reinforcement the children chose container A about 75% of the time, whereas under the high reinforcement condition they chose container A more than 75% of the time, thus maximizing their chances of being rewarded. In addition, other studies (Davies, 1965; Goldberg, 1966; Messick and Solley, 1957; Stevenson and Weir, 1959; Stevenson and Zigler, 1958; Yost, Siegel and Andrews, 1962) have found that the concept of probability is acquired earlier than hypothesized by Piaget.

In addition to the two choice probability tasks with children, there have been studies which introduced rewards for correct decisions and losses for errors. This procedure represents a more complex type of decision making. The general conclusion of these studies is that children, ages 5-11, maximize their rewards contingent on an event by increasing the frequency of the decision that the event will occur (Brackbill, Kappy, and Starr, 1962; Das and Panda, 1963; Offenback, 1964; Siegel and Andrews, 1962).

Piaget's failure to find probabilistic thinking in children earlier than the stage of concrete operations (7 years of age) has frequently been attributed to methodological problems in his search (Davies, 1965; Siegel and Andrews, 1962). Some of these problems include utilizing tasks that rely too much on verbal ability, are too complex, and which do not motivate the child with a reward system.

Description and Rationale

It is quite obvious that children at a very early age make decisions based on information which is probabilistic in nature. Indeed the decisions that children often make are quite complicated and involve the simultaneous

processing of information from multiple inputs. For example, let's say a child is making noise while playing and his mother says "stop making that noise." The child must now make a decision to either stop or continue making noise. Before this decision can be made the child must first decide if the mother really meant what she said. To do this the child must consider several dimensions of her message. The tenor and loudness of her voice, her facial expressions, and other aspects of the message must be considered. For the purpose of this illustration let me focus on just one of the dimensions of her message -- the loudness of her voice. Figure 1 may be used to clarify the discussion which is to follow. The X axis in figure 1 represents the loudness of the mother's voice in terms of decibels.

[Insert Figure 1 about here]

The point here is that the child has previously discovered that the loudness of the mother's voice is a critical variable to consider in deciding on the seriousness of her messages. Based on past experience the child has learned that when his mother says "stop making that noise" in a soft voice she usually doesn't mean it. When she screams out the message she usually does mean it. Unfortunately, there is a range of decibels within which the child is most uncertain, but nevertheless a decision must be made. In making this decision if the child is deciding rationally several variables will effect the decision outcome.

1. A Priori Probability - the child must have some idea of the odds in favor of the mother meaning what she says. For example if in the past the mother has not meant it 66% of the time and did mean it 33% of the time then the odds are 2 to 1 that she does not mean it.

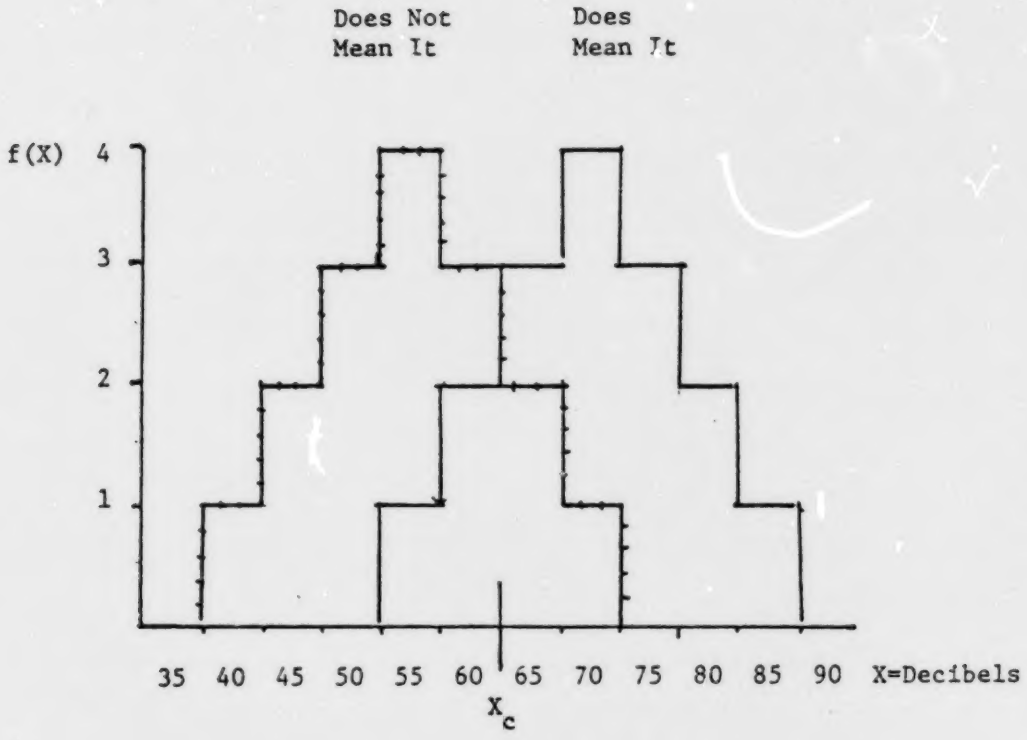


Figure 1. Hypothetical Distributions of the Loudness of the Mother's Message Expressed in Decibels.

If the child is making rational decisions based on an intuitive understanding of a priori probability, in this situation he should more often decide that she doesn't really mean it and continue making noise at play.

2. Conditional Probability - the child must know the number of times she meant it, or didn't mean it at each decibel level. This is represented by the height of the bars in figure 1. The distribution on the left reflects the number of times at each decibel level the mother in the past has said "stop making that noise" but didn't mean it. The distribution on the right reflects the number of times at each decibel level the mother in the past has said "stop making that noise" and meant it. For example, take a look at X = 65. The height of the bars indicate that the mother has said "stop making that noise" 5 times at this decibel level. Three times she did really mean it, but two times she did not. If the mother's message is at this decibel level, and if the child is making rational decisions based on an intuitive understanding of conditional probabilities then the child should most often decide that she meant it.

3. A Posteriori Probability - after the child takes into account a priori probability and conditional probability of her message he is now in a position to estimate the probability that she meant what she said or did not mean what she said. Even if the prior odds of her meaning it is 2 against 1, if the message were delivered at 80 decibels, the rational child will more likely decide that the mother really means to stop making noise. Another way of saying this is even though initially there is a greater chance that the mother doesn't

really mean what she says, the loudness of her message is so great as to lead the child to revise the a priori probability in the light of this new information.

This facet of probabilistic thinking has never been investigated with children although it has been researched in adults. A statistical theorem developed by the Reverend Thomas Bayes in 1736 allows one to actually determine these a posteriori probabilities. It is known that adults are very conservative in their estimates of a posteriori probabilities. One question this proposed research is attempting to answer is how does the performance of children compare with that of adults.

4. Utilities or Costs and Values - the child must also consider the cost of making the wrong decision and the value of making the right decision. If strong punishments are likely to follow if the child continues making noise when the mother means to stop, the rational child will not need so loud a message in order to make the decision to stop making noise.
5. Decision Criterion - information on the above four variables, a priori probability, conditional probability, a posteriori probability, and utility must be processed simultaneously if the child is to make nonrandom decisions under conditions of uncertainty. One way of conceptualizing this in terms of figure 1, is that the child must have some decision criterion, along the X axis (call it Xc). If the loudness of the message equals or exceeds Xc then the child decides the mother does not mean it. The rational child will change this decision criterion as a priori probabilities, conditional probabilities,

a posteriori probabilities, and utilities change. The Xc shown in figure 1 is the one which is optimal (i.e., most rational) when a priori probabilities and utilities are equal for both distributions. If the a priori probabilities are changed such that the left distribution is more probable, then Xc should move to the right. In other words, if the child is processing all the information correctly he should change his mind about the meaning of the mother's message. Indeed she would have to scream louder before the child will decide that she really means it. Other examples could be given requiring the most optimal decision criterion to be one to the left of that shown in figure 1.

As mentioned earlier, previous research has seldom specified what type of probability is being investigated, and has provided no systematic way of combining various types of probabilities and utilities in one experiment. The mathematical model proposed in this research was developed first in the engineering of sensing devices and later used extensively in the theory of signal detection in psychophysics and statistical Bayesian decision theory. Under the conditions specified by the researcher the model provides a precise mathematical measurement of the expected decision criterion. This mathematical model (which underlies the probability task to be used in this research) incorporates a priori probability, conditional probability, a posteriori probability and utility (all manipulated by the experimenter) in such a way as to determine the most "rational" decision under conditions of uncertainty. The details of the mathematical model are relatively complicated and will only be briefly discussed here.

Theory of Signal Detection

The TSD model will now be presented. For other sources with treatments at differing levels of complexity, consult: Combs, Davies and Tversky, 1970; Green and Swets (1966); Licklider (1959); Luce (1963) Swets (1964); Tanner and Main (1969).

TSD as a model is a mathematical similation of the human decision making process under conditions of uncertainty. The model combines features of (1) the engineering of sensing devices, and (2) statistical decision theory. When applied to human decision making the model provides a precise mathematical measure of (1) the upper limit, or ideal reception of ambiguous information, and (2) the optimal strategy for maximizing correct decisions or minimizing errors. The emphasis of the current study is on the second component of the model.

In a typical TSD experiment the task is to decide if an ambiguous stimulus arises from one of two distributions. For example, in an acoustical experiment the decision task is to decise of an auditory stimulus is noise (normally distributed, band limited white gaussian noise), or signal (some fixed signal added to the white gaussian noise).

The first component of the TSD model is concerned with how well the observer can perceive the difference between the two distributions. When the parameters of the noise and signal distributions are known (as with the current experiment), the ideal perception of the difference between the two distributions is equal to the standarized difference between the means of the two distributions (and is called d') given by the following formula:

$$d' = \frac{\mu_s - \mu_n}{\sigma_n}$$

One way of conceptualizing the TSD decision task is provided in Figure 2.

Insert Figure 2 About Here

The decision task is to decide whether an observation is signal or noise. In order to decide whether a signal is present, the observer must decide on some critical value of X (call it X_c), such that when $X \geq X_c$, the decision is that a signal occurred. When $X < X_c$, the decision is that noise occurred.

When the prior probabilities of the signal and noise distribution are equal, and the costs and values of decision outcomes are equal, the point X_c (where the distributions overlap) is the optimal cutoff point along the X axis which maximizes correct decisions and minimizes errors. However, if $P(S) > P(N)$ then this should create a tendency on the part of the decision maker to move the cutoff point X_c to the left, and conversely to the right if $P(S) < P(N)$. In addition to prior probabilities, the value of a correct decision and cost of an error will affect X_c. For example, if the value of detecting the signal is much greater than the value of detecting the noise then X_c will move to the left to increase valuable hits, and diminish costly false negatives. Conversely, if the value of detecting the noise is much greater than the value of detecting the signal then we will move X_c to the right, thereby increasing correct rejections and reducing costly false positives.

It can be shown that X_c is not the most useful criterion cutoff measure. Instead the liklihood ratio , $l(X)$, is used.

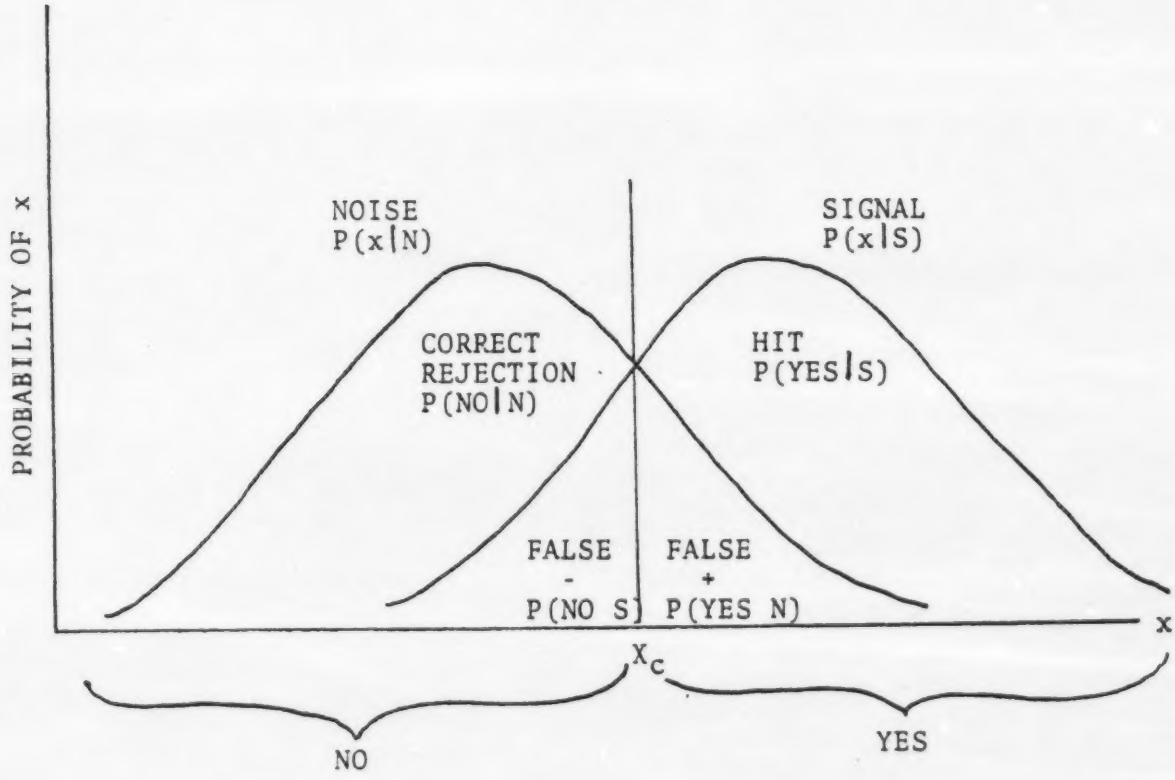


Figure 2. Noise and Signal Distributions in a Typical Signal Detection Task.

Liklihood ratios are the ratio of $P(X|S)/P(X|N)$, for each value of X. The conditional probabilities are indicated by the heights, or ordniates of the curve for each value of X, under each distribution. The criterion cutoff gives you the same decision as X_c. In fact, B is equal to the ratio of the ordinate of the signal distribution to the ordinate of the noise distribution at point X_c (i.e., B equals the l (X) at point X_c).

Given that the value function to be maximized is expected value, the most optimal B, which takes into account prior probabilities, and costs and value is:

$$B = \frac{P(N)\ VALUE(NO|N) - COST(YES|N)}{P(S)\ VALUE(YES|S) - COST(NO|S)}$$

When B is computed according to the above formula, it is the optimal criterion. The decision maker says YES (a signal occurred) when $l(X) \geq B$ (i.e., $X \geq X_c$), and says NO (a signal did not occur) when $l(X) < B$ (i.e., $X < X_c$).

As was stated earlier TSD provides the optimal strategy for maximizing correct decisions and minimizing errors. The optimal value of l (X), or B, is determined mathematically. It is based on the prior probabilities and utilities manipulated by the experimenter. The subject's B is computed from his actual performance. In an acoustical experiment the actual hit and false positive rate are entered into a standard normal table, and the ratio of the ordinates at X_c is B. Also, recall that TSD provides a mathematical measure of the upper limits, or maximum reception of ambiguous information. The ideal d' is manipulated by the experimenter, and is the standardized

difference between the means of the noise and signal distributions. The actual d' of the decision maker is determined empirically. Given the distributions are normal with variances equal to one, the differences between the standard score of X_c under the noise and signal distributions give you d', i.e., $d' = Z_N - Z_S$.

Looking again at Figure 2, it should be noted that no matter where the X_c is located along the X axis, the distance between the means is unaffected. Likewise, regardless of the distance between the means, the optimal cutoff which maximized expected value is given by B. Thus TSD provides two independent measures of these parameters. d' is a measure of the amount of information processed, whereas B is a measure of the ability to maximize expected value given the amount of processed information.

There are several advantages in using TSD in studying children's decision making in probabilistic environments. One advantage is that TSD allows the researcher to study independently two components of decision making often confounded in performance. These are (1) how much of the available information relevant to the decision is processed; and (2) how much of the processed information is utilized in making the decision. Particularly in young children it seems reasonable to want to separate these components. If a young child performs at chance level, we may want to know if the problem is a sensory one or an inability to utilize information that has been processed.

A second advantage of TSD is that the model includes a mathematically precise statement about the expected effect of values and prior probabilities on the decision criteria. This provides a standard against which the childs performance can be compared.

A third advantage of TSD is that it is a model that has been well researched with adults. Therefore, the performance of children within the Context of TSD can be compared to adults to get some idea of the developmental course of decision making under conditions of uncertainty.

The major purposes of the study were; (1) to provide evidence for the experimental hypothesis that the decision criteria of children within the study is age related, i.e., the older the child, the more his performance approximates optimal performance, and (2) to determine the appropriateness of TSD measures and methodology for childhood populations.

METHOD

SUBJECTS: Twenty-four subjects were selected for the study. Eight from a kindergarten and 8 from a college in Kanawha County, West Virginia.

Another 8 were from an elementary school in Putnam County, West Virginia. There were 8 subjects at each of 3 age levels (5, 11, and 18 years of age). These age ranges were selected based on the extant literature on probabilistic thinking in children, and a pilot study using three children and two adults.

In order to minimize individual differences in the subject selections, there was an attempt to control for sex and intelligence. Eight females of average intelligence were selected at each age level. Average intelligence was operationally defined as scores within 80-120 on the Peabody Picture Vocabulary Test (PPVT). Controls for I.Q. were included as there is minimal evidence that adult retardates perform less well than adult normals (Tanner and Main, 1969). Letters requesting parental permission for each kindergarten and elementary school child were obtained before the task was undertaken

MATERIALS: It was decided that auditory inputs would not be used as in the classic TSD acoustical experiment. In order to expedite matters and increase the availability of subjects, a task was developed that contained all of the elements of the TSD paradigm but was more appropriate for very young children. The task was more appropriate in that it: (1) was transported to the child rather than transporting the child to the elaborate and cumbersome equipment needed for the acoustical experiment; (2) involved inputs relevant to adults as well as young children; and (3) enabled the experimenter to interact with the child and thereby insure that the child remained motivated across trials.

Three sets of 72 (4" by 2½) cards were printed. Half of each set (i.e., 36) was used for the noise distribution, and the other half for the signal distribution. Both distributions consisted of a distribution of 3/4" wide black bars on the front of the cards. Each unit of increase in height of the bars was 3 mm. The height of the bars under the noise distribution ranged from 40 mm to 70 mm. The overlapping signal distribution ranged in height from 49 mm to 79 mm. The frequency of each bar height and associated conditional probabilities are given in Figure 3 and the hit and false positive rates age given in Table 1. On the back of the cards composing the signal distribution was a star, while there was a dot on the back of the cards composing the noise distribution. The child was required to decide whether a star or a dot was on the back of the card based on the height of the bar. Pennies were used to manipulate costs and values.

INSERT FIGURE 3 AND TABLE 1 ABOUT HERE

PROCEDURE: The subject was simultaneously shown two cards with the heights of the bars representing the means of the noise and signal distribution (55 mm and 64 mm respectively). The subject was then told:

> "We are going to play a game in which you must guess whether there is a star or a dot on the back of these cards. The only hint you will be given is this black bar on the front of the card. When the bar is tall there is usually a star on the back. On the other hand, when the bar is short there is usually a dot on the back. Now, I'm going to mix up these two

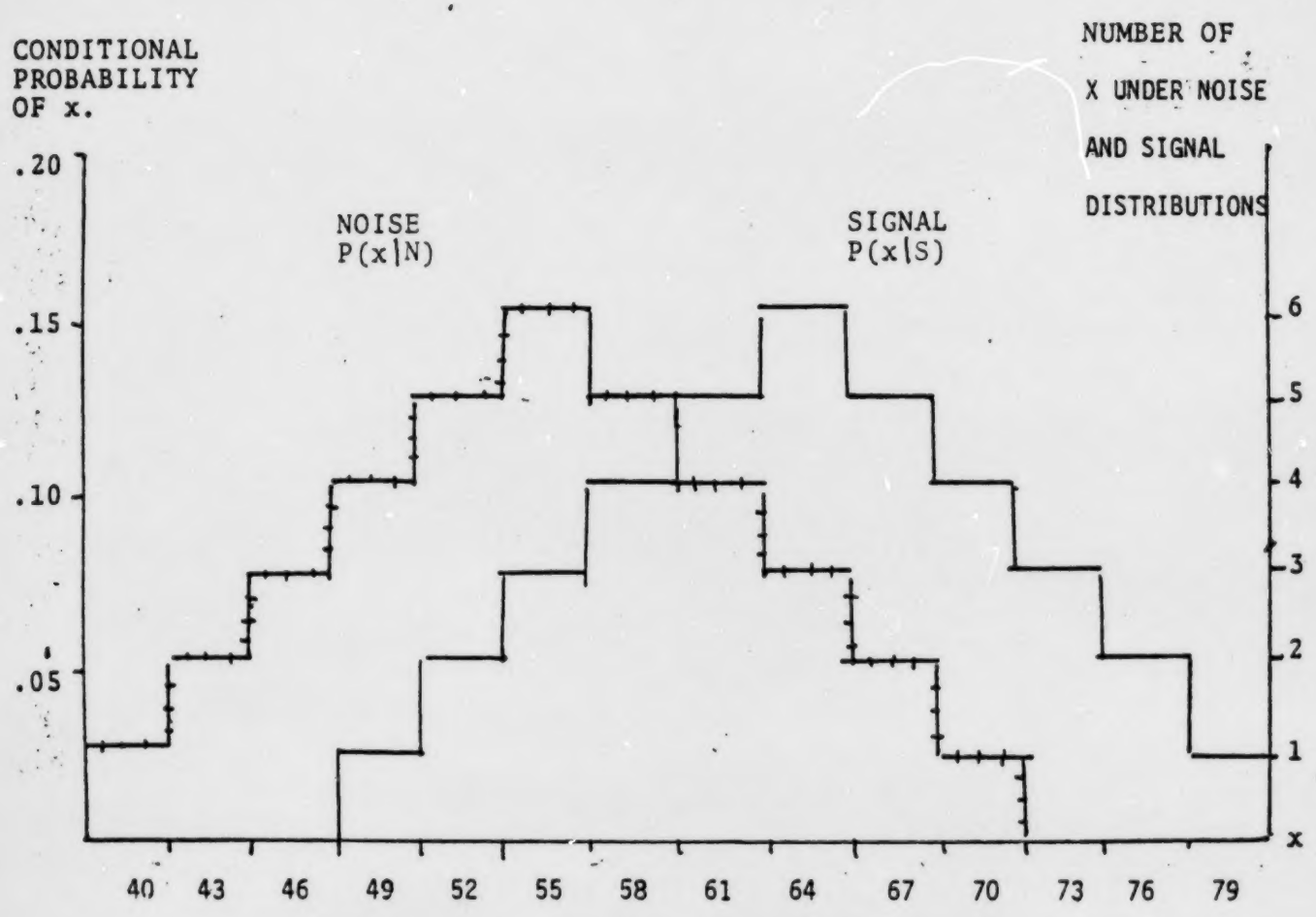


Figure 3. Distribution of Heights of Bars (in millimeters).

Table 1

Hit, False Positive Rate, and Likelihood Ratio of Bar Heights

Height of Bars in Millimeters	P(x|N)	False Positives P(yes|N)	P(x|S)	Hit Rate P(yes|S)	Likelihood Ratio P(x|S) / P(x|N)
40	.0277	.9718	.0000	1.0000	0.00
43	.0555	.9160	.0000	1.0000	0.00
46	.0833	.8330	.0000	1.0000	0.00
49	.1111	.7219	.0277	1.0000	0.25
52	.1388	.5831	.0555	.9718	0.40
55	.1667	.4164	.0833	.9160	0.50
58	.1388	.2776	.1111	.8330	0.80
61	.1111	.1665	.1388	.7219	1.25
64	.0833	.0832	.1667	.5831	2.00
67	.0555	.0277	.1388	.4164	2.50
70	.0277	.0000	.1111	.2776	4.00
73	.0000	.0000	.0833	.1665	∞
76	.0000	.0000	.0555	.0832	∞
79	.0000	.0000	.0277	.0277	∞

cards and you tell me which of these cards is more likely to have a star or a dot on the back. Base your decision on how tall or short the bar is on the front."

To insure discrimination between the means of the distribution, this procedure was repeated until the subjects got 3 consecutive guesses correct. In addition, 30 demonstration trials using all 72 cards were administered in each condition to give the subject practice with the task. The above instructions were repeated as many times as necessary until the experimenter felt the subjects understood the task. Also, throughout the experiment subjects were occasionally reminded (particularly the younger subjects) of the nature of the task. This was done because of findings in the pilot study. It was observed that subjects occasionally used strategies other than the height of the bar to make decisions.

The following additional instructions were given for each condition of the experiment.

EQUAL UTILITIES AND EQUAL PRIOR PROBABILITIES

"Please notice that I have the same number of cards with stars on the back as I have with dots on the back (36 signal and 36 noise cards). As you can see the two stacks of cards are the same size. Now watch me shuffle these two stacks together real well (well randomized). I'm going to show you these cards one at a time. Each time you look carefully at how tall the black bar is. Then decide whether or not a star or dot is on the back. Each time you are correct you will win one penny, and

each time you are incorrect you will loose one penny. Now lets do a few (30 practice trials) so you can see what I mean. I want you to win as many pennies as you can. The closer you pay attention and make decisions based on how tall or short the bar is the more pennies you will win."

UNEQUAL UTILITIES AND EQUAL PROBABILITIES

"Please notice that I have the same number of cards with stars on the back as I have with dots on the back. As you can see, the two stacks are the same size. Now watch me shuffle these two stacks together real well. I'm going to show you these cards one at a time. Each time you look carefully at how tall the bar is. Then decide whether a star or a dot is on the back. The cards with a dot on the back are worth one penny,whereas the cards with a star on the back are worth 3 pennies. When you get the cards with a dot on the back right you will win one penny, whereas when you get them wrong you will loose one penny. When you get the cards with a star on the back right you will win 3 pennies, whereas when you get them wrong you will loose 3 pennies. Remember, the cards with a star on the back are worth 3 pennies. Now let's do a few so you can see what I mean. I want you to win as many pennies as you can. The closer you pay attention and make decisions based on how tall or short the bar is the more pennies you will win.

EQUAL UTILITIES AND UNEQUAL PRIOR PROBABILITIES

"Please notice that I have three times as many cards with stars on the back as I do cards with dots on the back. As you can see the stack of cards with stars on the back are three times as tall as the stack of cards with dots on the back. Now watch me shuffle these two stacks together real well. I'm going to show you these cards one at a time. Each time, you look carefully at how tall the bar is. Then decide whether a star or a dot is on the back. Each time you are correct you will win one penny, and each time you are incorrect you will loose one penny. Now lets do a few so you can see what I mean. I want you to win as many pennies as you can. The closer you pay attention and make decisions based on how tall or short the bar is the more pennies you will win."

The subject was "spotted" ten pennies to begin the game. The experimenter held the cards before the subject, and once a decision was made as to whether a star or a dot was on the back the card was turned over and placed in one of two piles (the child received or lost a penny after each trial). One pile consisted of cards the child decided had a dot on the back and the other those he/she decided contained a star. At the end of 72 trials, the hits and false positives were counted. Before and after each series of trials the deck of cards was well shuffled to insure randomization. There were 288 trials per condition (four series of 72 trials) for each subject.

It was found in the pilot study that the task become tedious and fatiguing when each subject was subjected to each condition (1152 trials). This was true even when frequent, or prolonged breaks were used. Therefore, in order to get the best effort from each subject a between subject design was used, such that with half of the subjects only criterion levels of value were manipulated, and with the other half only criterion levels of prior probability were manipulated. This required only 576 trials per subject.

The independent variables in the study were: (1) age of the child; and (2) value and prior probabilities of the signal and noise distribution. The dependent variable was the decision criterion, B, based on the subjects actual performance. Although measures of d' were included in the study, they were not relevant to the major emphasis of the research, i.e., the application of the decision making component of the TSD model to a childhood population.

The two levels of the value manipulation are presented in the following payoff matrices along with their expected B value.

1)

	YES (S)	NO (S)
S	1¢	-1¢
N	-1¢	1¢

B = 1.00

2)

	YES (S)	NO (S)
S	3¢	-3¢
N	-1¢	1¢

B = .33

The two levels of the prior probability manipulation are presented in the following ratios along with their expected B value.

1) $\frac{P\ (N) = .5}{P\ (S) = .5}$ $\quad B = 1.00$

2) $\frac{P\ (N) = .25}{P\ (S) = .75}$ $\quad B = .33$

These two manipulations (i.e., costs and values along with prior probabilities) were set orthogonal to each other such that the second payoff matrix was used only under conditions of the first prior probability ratio, and the second prior probability ratio was used only under conditions of the first payoff matrix. Subjects were assigned randomly to the experimental conditions, and the order of the B level was counterbalanced. For example, the order of the B level for the first two day care children was B = 1.00, then B = .33. For the second two children the order was B = .33, then B = 1.00.

The level of significance established for acceptance of results was set at .01.

Results

The results of the study may be divided into two parts, each related to the two major purposes of the study. These major purposes were to determine whether: (1) children change their decision criteria as a function of the age levels used in the study; and (2) TSD measures and methodologies were appropriate for childhood populations.

(1) Decision Criteria (B) as a Function of Age

The experimental hypothesis of the study was that the older the child, the more his performance approximates optimal performance. To test this hypothesis, a 3 X 2 X 2 analysis of variance was used with repeated measures on the third factor (Table 2). Evidence for the assumption of homogeneous treatment populations was provided by the F max ratio which was non-significant at the .01 level (f max (1,18) = 7.78, $P > .01$). The first independent variable was age with three levels; day care, elementary school, and college. Contrary to expectation, there was no difference in how much the subjects changed their decision criteria across age levels. Therefore these results of the study do not support the experimental hypothesis.

Insert Table 2 About Here

(2) Appropriateness of TSD Measures (i.e., B and d') and Methodologies for Childhood Populations

Since past TSD research has not included childhood populations, it was not known whether young children made decisions as predicted by the model. Therefore, this study was viewed as an initial attempt at determining the adequacy of TSD in predicting children's decisions under conditions of uncertainty. The two TSD measures that this study investigated for childhood populations were the decision criterion, B, and the measure of perceptual discrimination, d'. Results relevant to each of these will be discussed separately.

(2.a.) Appropriateness of Decision Criterion (B) Measure for Childhood Populations

It was reasoned that if the subject's decision criteria conformed to

Table 2

Analyses of Variance:

Measure of Criterion Change (B)

Source	df	MS	F
Between Subjects	23		
Age Level	2	.04432	1.53
Probability - Value	1	.35719	12.35**
Age X Prob. - Value	2	.01221	.42
$Error_b$	18	.02893	
Within Subjects	24		
Criterion Level	1	.75543	312.16***
Criterion X Age	2	.00714	2.95
Criterion X Prob.- Value	1	.02064	8.53**
Criterion X Age X P.- V.	2	.00063	.26
$Error_w$	18	.00242	
Total	47		

**$p < .01$

***$p < .001$

the assumptions of the model, then the actual B changes should be characterized by two tendencies: (1) they should change more than would be expected by chance, and (2) they should change in the direction of optimal B.

To test for this first tendency, the expected criterion level of B was included as the third variable in the analysis of variance presented in Table 2. This variable involved within group, or repeated measures as the same subjects were used at both expected criterion levels (i.e., B = 1.00 to B = .33, or B = .33 to B = 1.00). The main effect of the criterion level was highly significant (F (1,18) = 372.01, $P < .0001$) confirming that the subjects changed their decision criteria more than would be expected by chance.

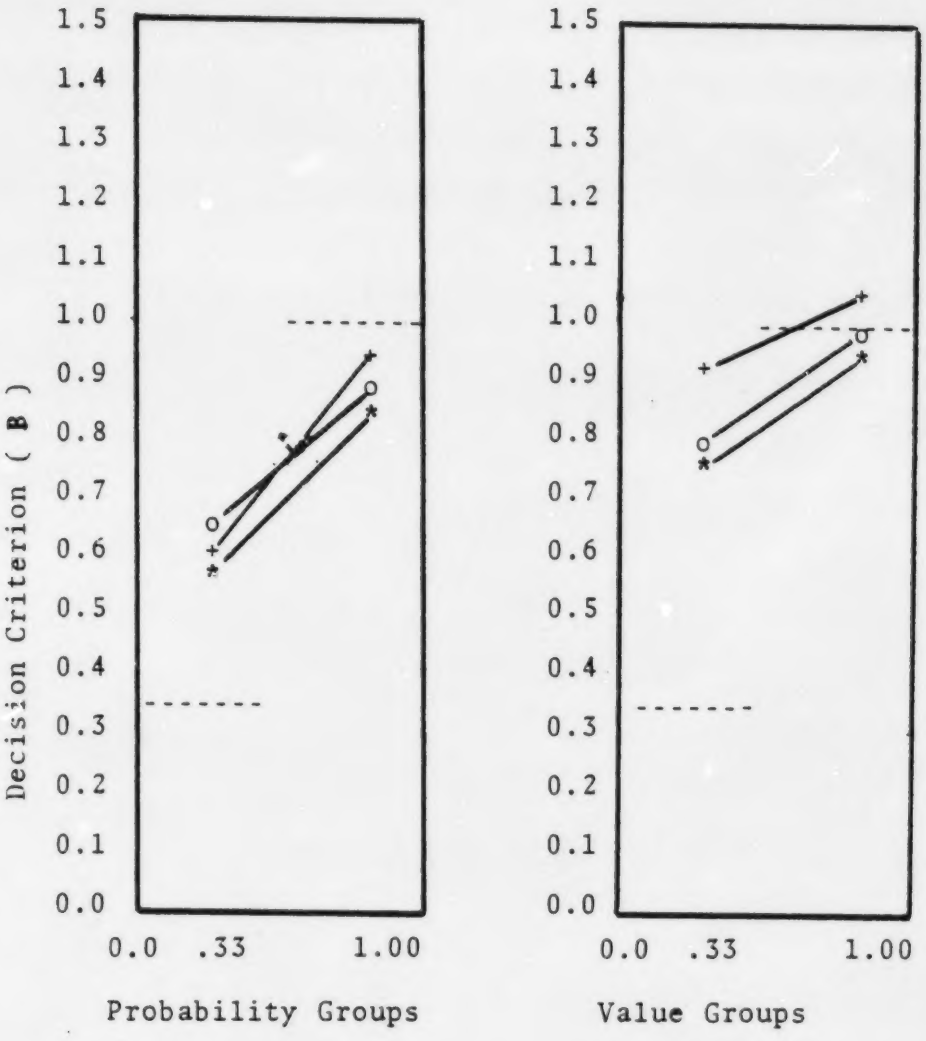
In order to determine if the decision criteria changed in the direction of optimal B, the actual B changes were graphed in Figure 3.

Insert Figure 4 About Here

In figure 4, the mean changes in B for each experiment condition are plotted against the optimal criterion change. An inspection of figure 3 indicates a general ordinal correspondence between the actual and optimal criteria.

The results of the study therefore indicate that the decision criteria of children do operate according to the assumptions of the TSD model. Therefore, the measure of criterion change, B, appears to be an appropriate one for investigating probabilistic thinking in young children.

There were several significant outcomes in the analysis of variance presented in Table 2 that were not predicted under the TSD model. For example,



Note. Horizontal lines indicate optimal criteria

Figure 4. Mean changes in B for each experimental condition.

the main effect of the probability - value independent variable was significant (F (1,18) = 12.38, $P < .01$). An inspection of the mean B changes in figure 3 indicates that across all age levels the actual decision criteria tended to be lower when prior probabilities were manipulated than when values were changed. This suggests that there was more of a bias toward the signal cards under the prior probability condition. Under the TSD model the bias, or preference for the signal should be the same regardless of whether values or prior probabilities are being manipulated. In other words, the actual criteria should be the same when the signal is worth 3 times more as when there are 3 times as many signals.

Another outcome which was not predicted under the TSD model was the significant interaction of the criterion levelwith the probability - value independent variable (F (1,18) = 8.53, $P < .01$). Collapsing across age levels this was a 2 X 2 interaction with six possible cell comparisons. A procedure suggested by Winer (1971, pp. 565-567) indicated that all six individual cell by cell comparisons were significant at the .01 level. Each of the six cell by cell comparisons will be discussed separately.

The significant difference between the criterion levels when prior probability (F (1,18) = 211.12, $P < .01$) and value (F (1,18) = 108.8, $P < .01$) were manipulated was interpreted as meaning that the main effect of the criterion level (reported above) was true for both levels of the probability - value independent variable. Therefore, these two significant cell comparisons were expected under the TSD model.

However, the significant differences between the probability and value

manipulation when B = 1.00 (F (1,18) = 42.65) $P < .01$) and when B = .33 (F (1,18) = 111.35, $P < .01$) were not expected under the TSD model. These significant cell differences indicate that the main effect of the probability - value independent variable (reported above) was true for both criterion levels. Therefore, this suggests that the greater bias toward the signal cards when prior probability was manipulated was true for both criterion levels of B.

The last two cell comparisons were: (1) the value manipulation when B = 1.00 with the probability manipulation when B = .33 (F (1,18) = 444.57, $P < .01$) and (2) the value manipulation when B = .33, with the probability manipulation when B = 1.00 (F (1,18) = 15.42, $P < .01$). The significant differences between these cell comparisons were also expected under the TSD model.

To summarize the anomalous outcomes in the study it was found that the subject's bias toward the signal cards was greater when prior probabilities were manipulated than when values were changed. Furthermore, the subjects were more biased toward the signal cards at both criterion levels of B. Although it is possible that these outcomes represent actual performance characteristics of the subjects, the researcher suspects that a methodological procedure employed in the study reasonable explains the anomalous outcomes. This explanation will be reserved for the discussion section.

(2.b.) Appropriateness of Perceptual Discrimination Measure d' for Childhood Populations

Although less relevant to this thesis topic of probabilistic thinking in children, the TSD measure of perceptual discrimination (d') was also

investigated for each subject. If the subjects were operating under the TSD model, it was expected that none of the experimental manipulations should have affected d'. All of the experimental manipulations employed in the study should affect only the decision criteria, and not how well the subject discriminates between the noise and signal. Since the B and d' changes are independent parameters, the d's in the study should only differ by chance.

To test this expectation, a 3 X 2 X 2 analysis of variance was employed with repeated measures on the third factor. This is presented in Table 3. It should be noted that the design is identical to the analysis

Insert Table 3

of variance used with the B changes, except d' is now the dependent variable. There were no outcomes in the analysis which were significant at the .01 level. The main effect of the criterion level, however, was significant at the .05 level. (F (1,18) = 6.66, $P < .05$). Since the significance level was set at .01 the outcome does not meet the criterion established for the acceptance of results.

The results of the study therefore indicate that the experimental manipulations did not significantly effect d'. The results of the study suggest that the measure of perceptual discrimination does operate according to the assumptions of the TSD model for childhood populations.

Table 3

Analyses of Variance:

Measure of Perceptual Discrimination (d')

Source	df	MS	F
Between Subjects	23		
Age Level	2	.26104	2.27
Probability - Value	1	.09929	.86
Age X Prob. - Value	2	.02937	.23
$Error_b$	18	.11498	
Within Subjects	24		
Criterion Level	1	.20703	6.66*
Criterion X Age	2	.03805	1.22
Criterion X Prob.- Value	1	.09242	2.97
Criterion X Age X P.-V.	2	.02711	.87
$Error_w$	18	.03109	
Total	47		

*$p < .05$

Discussion

One major purpose of the current study was to determine if probabilistic thinking within a TSD framework is age related. The experimental hypothesis of the study was that the older the child the more his performance approximates optimal performance. However, the main effect of the age variable was nonsignificant, and the age interactions were nonsignificant in the analysis of variance. Therefore, the results of the study do not support the experimental hypothesis that the decision criteria of children within the context of the TSD paradigm is age related. These results have implications for the general literature on probabilistic thinking in children.

The results of the study do not support Piaget's theory that the onset of probabilistic thinking in children occurs between 7 - 10 years of age. Rather the results suggest that children as early as 4 to 5 years of age solve probabilistic tasks as well as do adults. This is consistent with much of the past American research (Davies, 1958; Goldberg, 1966; Messick and Solley, 1957; Siegel and Andrews, 1962; Stevenson and Weir, 1959; Stevenson and Zigler, 1958; and Yost, 1962) even though different probability tasks were used.

In addition to addressing itself to the literature on probabilistic thinking in children, another purpose of the study was to determine the appropriateness of TSD measures and methodologies for childhood populations. It was found that young children do operate as predicted by the TSD given the conditions specified by the model. Therefore, TSD

might serve as a useful research paradigm for further studies investigating probabilistic thinking in children. Also, a task may be developed which utilize TSD measures and methodologies yet is appropriate for very young children as well as adults.

However, one significant outcome in the study was inconsistent with findings utilizing other TSD tasks (Swets, 1959). The anomalous outcome was the significant difference between the probability and value manipulation for both criterion levels of B. It is the researchers belief that this outcome is a probable artifact of a procedure employed in the task rather than representing actual characteristics of the subjects performance.

Although subjects were assigned randomly to groups, and the order of the criterion level of B was counterbalanced within groups, the order of the groups at each age level was not counterbalanced. At each age level values were manipulated first followed by the probability manipulation. By the time the first 4 subjects in the value group were finished, the may have communicated to other students that the cards with a star on the back were more important, creating a bias favoring the signal cards in the probability groups. In future studies using the current task this situation may be corrected by using more subjects at each age level and counterbalancing the order of the probability and value groups, or randomly alternating the use of the dot and the star as the signal. In the current study, only the star was used as a signal.

In addition to methodological considerations in the study, the researcher feels that a discussion is needed as to whether the data in the experiment meet the assumptions underlying the use of the analysis of variance.

Two major assumptions of the analysis of variance are (1) that the population variances are homogeneous, and (2) that the treatment population is normal. The nonsignificance of the F max ratio may be interpreted as providing evidence that the assumption of homogeniety is warranted.

Concerning the normality assumption, although it has not been well tested it is assumed in the TSD literature that the population of B is gaussian in form (Egan, 1977). Futhermore, B may be assumed to be normally distributed for two overlapping triangular distributions (as in the current experiment) as well as for two overlapping normal distributions (Egan, 1977).

Although B is assumed to be normally distributed, since little research has been done to confirm this, the researcher is less confident that the data meet the assumption of normality. The possible violation of this assumption, however, is partially compensated for by the fact that a conservative alpha level of .01 was used in the study and that the analysis of variance is robust with respect to moderate departures from normality.

REFERENCES

Bell, Donald W., and Nixon, James. Reliability of Ratings in an Auditory Detection Experiment. Journal of the Acoustical Society of America, 1971, 49, 435-439.

Brackbill, Y., Kappy, M.S., and Starr, R.H. Magnitude of Reward and Probability Learning. Journal of Experimental Psychology, 1962, 63, 32-35.

Coombs, Clyde H., Dawes, Robyn M., and Tversky, Amos. Mathematical Psychology: An Elementary Introduction. Englewood Cliffs, New Jersey: Prentice - Hall, Inc., 1970, p. 165-201

Crandall, Vaughn J., Soloman, Dan., and Kellaway, Richard. A Comparison of the Patterned and Nonpatterned Probability Learning of Adolescents and Early Grade School Age Children. The Journal of Genetic Psychology, 1961, 99, 29-39

Das, J. P., and Panda, K.C. Two choice Learning in Children and Adolescents Under Reward and Non-reward Conditions. Journal of General Psychology. 1963, 68, 203-211.

Davies, Carolyn M. Development of the Probability Concept in Children. Child Development, 1965, 36, 779-788

Egan, James, Personal Communication, August, 1977

Egan, James, Signal Detection Theory and ROC Analysis. New York, NY: Academic Press, 1975

Fishbein, Efraim., Pampu, Ileana., and Manzat, Jon. Comparison of Ratios and the Chance Concept in Children. Child Development, 1970, 41, 377 - 389.

Fox, W. C. Signal Detectability: A Unified Description of Statistical Methods Employing Fixed and Sequential Observation Processes, The University of Michigan: Electronics Defense Group, 1953, Technical Report No. 19.

Goldberg, S. Probability Judgments of Pre-school Children: Task Conditions and Performance. Child Development, 1966, 37, 157 - 168.

Green, D. M., and Swets, John A. Signal Detection Theory and Psychophysics. New York: Wiley, 1966.

Gruen, Gerald., Ottinger, Donald, and Zigler, Edward. Level of Aspiration and the Probability Learning of Middle - and Lower - Class Children. Developmental Psychology, 1970, 3, 133 - 142.

Hoeman, Harry W., and Ross, Bruce M. Children's Understanding of Probability Concepts. Child Development, 1971, 42, 221 - 236.

Ingham, J. G. Individual Differences in Signal Detection. Acta Psychologica, 1970, 34, 39 - 50.

Katz, Leonard. Monetary Incentive and Range of Payoffs as Determiners of Risk Taking. Journal of Experimental Psychology, 1962, 64, 541 - 544.

Lathrop, Richard G. Developmental Factors in the Perception of Variability. Perceptual and Motor Skills, 1969, 29, 63 - 70.

Lewis, M. Probability Learning in Young Children: The Binary Choice Paradigm. Journal of Genetic Psychology, 108, 43 - 38, 1966.

Licklider, J. C. Three Aduitory Theories. In S. Koch (ed.). Psychology: A Study of a Science, Volume I. New York: McGraw Hill, 1959.

Luce, R. Duncan., and Green, David M. Detection of Auditory Signals Presented at Random Times. Perception and Psychophysics, 1970, 7, 1 - 14.

Luce, R. D., A Threshold Theory for Simple Detection Experiments. Psychological Review, 1963, 70, 61 - 79.

McGinnes, James H., and Berg, Norman L. Effects of Prior Gains on Risk - Taking Behavior by Children. Psychological Reports, 1973, 32, 543 - 549.

Messick, Samuel J., and Solley, Charles M. Probability Learning in Children: Some Exploratory Studies. The Journal of Genetic Psychology, 1957, 90, 23 - 32.

Offenback, Stuart. Studies of Children's Probability Learning Behavior: I. Effects of Reward and Punishment of Two Levels. Child Development, 1964, 35, 709 - 715.

Offenbach, Stuart. Studies of Children's Probability Learning Behavior: V. Sex Differences at Two (2) Age Levels. Journal of Genetic Psychology, 1968, 113, 95 - 100.

Offenbach, Stuart, Studies of Children's Probability Learning Behavior: VIII. Effect of Punishment at Two (2) Grade Levels. Psychonomic Science, 1971, 23, 407 - 408.

Piaget, Jean, and Inhelder, Barbel. La Genesis de L'idee de Hasord Chez Lienfant. Paris: Press University de France, 1951.

Piaget, Jean. The Childs Conception of the World. Patterson, New Jersey: Littlefield, Adams and Company., 1960.

Piaget, Jean. The Psychology of Intelligence, Patterson, New Jersey: Littlefield, Adams and Company., 1960.

Piaget, Jean. Une Experience sur la Psychologie du Hosard Chez L'enfant: le Tirage au Mart des Couples. Acta Psychologica, 1950, 7, 323 - 336.

Rosenfield, M., Copeland, B., and Suydam, Mary M. The Effects of Contingent Gains and Losses on Children's Risk Taking Behavior. Child Development, 1969, 40, 245 - 254.

Schmitt, Samuel A. Measuring Uncertainty: An Elementary Introduction to Bayesian Statistics. Reading, Massachusetts: Addison - Wesley, 1969, p. 330 - 337.

Siegel, Linda S. The Development of the Ability to Process Information. Journal of Experimental Child Psychology, 1968, 6, 368 - 383.

Siegel, S., and Andrews, J. M. Magnitude of Reinforcement and Choice Behavior in Children. Journal of Experimental Psychology, 1962, 63, 337 - 341.

Slovic, Paul. Risk Taking in Children: Age and Sex Differences. Child Development, 1966, 37, 169-176.

Stevenson, H.W., and Weir, M.W. Variables Affecting Children's Performance in a Probability Learning Task. Journal of Experimental Psychology, 1959, 57, 403-412.

Stevenson, H.W., and Zigler, E.F. Probability Learning in Children. Journal of Experimental Psychology, 1958, 56, 185-192.

Sullivan, Frank J., and Ross, Bluce M. What is Learning in Probability Learning. Developmental Psychology, 1970, 2, 58-65.

Swets, John A. (ed.). Signal Detection and Recognition of Human Observers: Contemporary Readings. New York: Wiley, 1964.

Swets, John A. A Brief Introduction to Decision Theory. Massachusetts Institute of Technology, Unpublished Manuscript.

Swets, J. A. Indices of Signal Detectability Obtained With Various Psychometric Procedures. Journal of the Acoustical Society of America, 1959, 31, 511-13.

Swets, J. A. Is there A Sensory Threshold?, Science, 1961, 134, 168-77.

Swets, John A., Tanner, Wilson P., and Birdsall, Theodore G. Decision Processes in Perception. Technical Report No. 40, Issued by The Electronic Defense Group of the University of Michigan, 1955.

Tanner, Wilson P. A Dice Game Illustration of Decision Theory, University of Michigan, Unpublished Manuscript.

Tanner, Wilson P. and Main, Dana B. Decision Criteria in Endogenous Mentally Retarded Adults, University of Michigan, 1969, Unpublished Manuscript.

Weir, Morton, W. The Effects of Age and Instructions on Child Probability Learning. Child Development, 1962, 33, 729-735.

Winer, B. J. Statistical Principles in Experimental Design, New York: McGraw-Hill, 1971.

Yost, Patricia A., Siegel, Alberta E., and Andrews, Julia M. Non-verbal Probability Judgements by Young Children. Child Development, 1962 33, 769-780.